

DERWENT-ACC-NO: 2003-426293

DERWENT-WEEK: 200340

COPYRIGHT 1999 DERWENT INFORMATION LTD

TITLE: **Text mining** method for electronic information stored in personal computer, involves dividing process document set **extracted** from pre-registered document set, and **extracting lexicons** corresponding to each divided document set

PATENT-ASSIGNEE: HITACHI LTD[HITA]

PRIORITY-DATA: 2001JP-0341474 (November 7, 2001)

PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES
MAIN-IPC JP 2003141134 A	May 16, 2003	N/A	015 G06F 017/30

APPLICATION-DATA:

PUB-NO	APPL-DESCRIPTOR	APPL-NO	APPL-DATE
JP2003141134A	N/A	2001JP-0341474	November 7, 2001

INT-CL (IPC): G06F017/30

ABSTRACTED-PUB-NO: JP2003141134A

BASIC-ABSTRACT:

NOVELTY - A process document set **extracted** from a pre-registered document set, is divided using a division standard. **Lexicons** corresponding to each divided document set, are **extracted from which the lexicon** having appearance deviation larger than a preset appearance deviation is taken as the **text mining** result.

DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for text mining processing apparatus.

USE - For mining text of electronic information stored in personal computers, word processors.

ADVANTAGE - Efficiently assists the text mining process, thereby providing user desired text mining result having resolution.

DESCRIPTION OF DRAWING(S) - The figure shows the flowchart illustrating the text mining method of electronic information stored in computer systems. (Drawing includes non-English language text).

CHOSEN-DRAWING: Dwg.2/15

TITLE-TERMS: TEXT MINE METHOD ELECTRONIC INFORMATION STORAGE PERSON COMPUTER

DIVIDE PROCESS DOCUMENT SET EXTRACT PRE REGISTER
DOCUMENT SET
EXTRACT CORRESPOND DIVIDE DOCUMENT SET

DERWENT-CLASS: T01

EPI-CODES: T01-J05B;

SECONDARY-ACC-NO:

Non-CPI Secondary Accession Numbers: N2003-340611